

Opinnäytetyö (AMK)

Tietojenkäsittely

Tietoliikenne

2010

Alexi Niiranen

VSSHP:n yhteinen hakemistojako

– uusimisprojektin selvitystyö



TURUN AMMATTIKORKEAKOULU
TURKU UNIVERSITY OF APPLIED SCIENCES

OPINNÄYTETYÖ (AMK) | TIIVISTELMÄ

TURUN AMMATTIKORKEAKOULU

Tietojenkäsittely | Tietoliikenne

Joulukuu 2010 | 25 sivua

Ohjaaja Esko Vainikka

Alexi Niiranen

VSSHP:n yhteinen hakemistojako - uusimisprojektin selvitystyö

Opinnäytetyön aiheena on Varsinais-Suomen sairaanhoitopiirin yhteisen hakemistojakon kartoitus. Työ tehtiin toimeksiantona Medbit Oy:lle. Työssä kerrotaan tekniikoista, jotka ovat hakemistojakon toteutuksen taustalla, tutkimuksen ongelmista, tutkimustulosten esittämisestä ja luokittelusta, hakemistojakon vaikutuksesta työntekijään ja ehdotuksia uusimisprojektin toteutukseen. Tutkimuksen tavoitteena oli selvittää hakemistojakon todellinen käyttö ja käyttäjät.

Työn teoriaosuuden pääpaino on palvelinklustereissa ja Distributed File Systemeissä. Klusterit ovat yhteenliitettyjä tietokoneita, jotka käyttäjä näkee yhtenä resurssina. Distributed File System, eli DFS, on teknologia, jolla jaetut hakemistot yhdistetään nimiavaruuksiksi. Nimiavaruus on virtuaalinen näkymä jaetuista hakemistoista. Teoriaosuutta havainnollistetaan kuvilla.

Tutkimuksen suurin ongelma oli hakemistojakon koko. Tutkimustulosten luokittelussa esiintyi myös ongelmia. Tulokset esitetään Excel-taulukoissa.

ASIASANAT:

klusterit, Distributed File System, DFS, hajautettu tiedostojärjestelmä, hakemistojako, Varsinais-Suomen sairaanhoitopiiri, VSSHP

BACHELOR'S THESIS | ABSTRACT

UNIVERSITY OF APPLIED SCIENCES

Business Information Technology | Data communications

December 2010 | 25 pages

Instructor Esko Vainikka

Alexi Niiranen

The Common Shared Directory of the Hospital District of Southwest Finland - research for rebuilding project

The subject of this thesis is the survey of the common shared directory of the Hospital District of Southwest Finland. The thesis was made as an assignment for Medbit Oy. The written part of the thesis consists of the theory about the technology behind the implementation of the directory, the description of the problems that occurred during the research, the presentation and classification of the results of the research, the effects of the directory to the employees and suggestions to the reforming of the directory. The goal of the research was to define the actual use and users of the directory.

The major part of the theory is about server clusters and Distributed File Systems. The clusters are computers that are connected and seen as one resource. Distributed File System (DFS) is a technology that allows shared directories to be connected as namespaces. A namespace is a virtual view of the shared directories.

The biggest problem of the research was the size of the directory. There were also problems in the classification of the research results. The results are presented in Excel documents.

KEYWORDS:

clusters, Distributed File System, DFS, shared directory, Hospital District of Southwest Finland

SISÄLTÖ

1	JOHDANTO	5
2	KLUSTERI	6
2.1	Klusteripalvelu	8
2.2	Resurssimonitori	8
2.3	Loogiset komponentit	9
2.4	Fyysiset komponentit	9
2.5	Failover ja failback	11
3	DISTRIBUTED FILE SYSTEM	11
3.1	Termit	12
3.2	Stand-alone ja domain-pohjainen DFS	13
3.3	DFS-replikointi	14
3.4	DFS:n rajoitukset	15
4	PÄÄSYNVALVONTALISTAT	16
5	SKRIPTIT	16
6	TUTKIMUKSEN ONGELMAT	17
7	HAKEMISTON ESITYS	18
8	Y-ASEMAN VAIKUTUS TYÖNTEKIJÄÄN	19
9	TIEDON LUOKITTELU	19
10	HAKEMISTON UUDELLEENRAKENNUS	22
11	JOHTOPÄÄTÖKSET	24
	LÄHTEET	25

KUVAT

Kuva 1.	Kuvakaappaus Excel-taulukosta, jossa luotiin skriptejä.	17
Kuva 2.	Esimerkki taulukosta, hakemiston sisältö.	18
Kuva 3.	Esimerkki taulukosta, hakemiston oikeudet.	18

KUVIOT

Kuvio 1.	Esimerkki klusteroidusta ratkaisusta (Shimonski 2003, 5).	7
Kuvio 2.	Esimerkki Active/Active klusteroinnista (Shimonski 2003, 14).	7
Kuvio 3.	Esimerkki Active/Passive klusteroinnista (Shimonski 2003, 13).	8
Kuvio 4.	Klusterin komponentteja.	10
Kuvio 5.	Failover ja failback (Shimonski 2003, 11).	11
Kuvio 6.	Prosessi tiedon etsimiseen DFS:stä clientin näkökulmasta (Boswell 2002).	12
Kuvio 7.	DFS nimiavaruus (Microsoft 2010).	13
Kuvio 8.	Stand-alone ja domain-pohjainen DFS (Microsoft 2003a).	14
Kuvio 9.	DFS-replikointi (Microsoft 2010).	14
Kuvio 10.	VSSHP:n DFS:n rakenne.	15
Kuvio 11.	Luokiteltu sisältö TYKS-kansiossa.	21

Kuvio 12. Luokiteltu sisältö muissa kansioissa.

21

1 Johdanto

Organisaatioilla voi olla tietoa, joka pitää olla kaikkien jäsenten nähtävillä tai käsiteltävissä. Yksi tapa toteuttaa tämä on yhteinen hakemistojako. Hakemistossa voi olla myös alihakemistoja, joihin vain osalla organisaatiosta on pääsy. Opinnäytetyön aiheena on Varsinais-Suomen sairaanhoitopiirin (VSSHP) yhteisen hakemistojaon (Y-asema) kartoitus. VSSHP:n tietohallintopalvelut on ulkoistettu Medbit Oy:lle, joka on opinnäytetyön toimeksiantaja.

VSSHP:n toimialueella on usea tuhat käyttäjää, mikä asettaa myös yhteiselle hakemistolle monia vaatimuksia. VSSHP:ssä on myös suuri määrä organisaatioita, joilla voi olla lukuisia aliorganisaatioita ja näillä vastaavasti aliorganisaatiota (sairaalat -> klinikat -> osastot), joilla on erilaisia tarpeita. Yksi vaatimus on korkea saatavuus (high availability), joka tässä tapauksessa on toteutettu klusteroinnilla. Hakemiston saatavuutta on lisätty myös Distributed File System -tekniikalla. Koska useassa hakemistossa saattaa olla potilas- tai tutkimustietoa tai muuta arkaluonteista tietoa, täytyy myös hakemistoihin pääsyä rajoittaa.

Tutkimus toteutettiin pääosin "brute force" -tekniikalla käymällä hakemisto kerrallaan läpi ja kirjaamalla, mitä se sisälsi. Tiedon moninaisuuden vuoksi tiedon keruuta ei voinut automatisoida kokonaan. Myös joidenkin hakemistojen fyysinen koko hankaloitti skriptien tai sovellusten käyttöä kartoituksessa. Skriptien avulla saatiin kartoitettua kuitenkin hakemistojen puu-rakenne ja pääsynvalvontalistat (ACL).

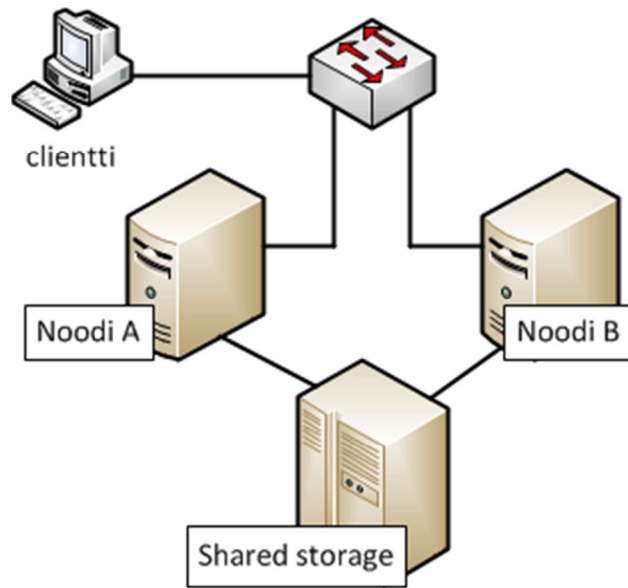
Tutkimuksen kohteena olevan hakemistojaon ongelmana on vapautunut rakenne ja oikean toiminnan hahmottaminen. Tutkimuksen tavoitteena oli selvittää hakemistojaon todellinen käyttö ja käyttäjät, mahdollisen uudelleen rakentamisen ja organisoinnin kannalta. Työssä käsiteltyä teoriaa tarkastellaan Microsoft Windows Server -tuoteperheen näkökulmasta ja palveluiden toteutukseen voi olla olemassa myös toisenlaisia ratkaisuja.

2 Klusteri

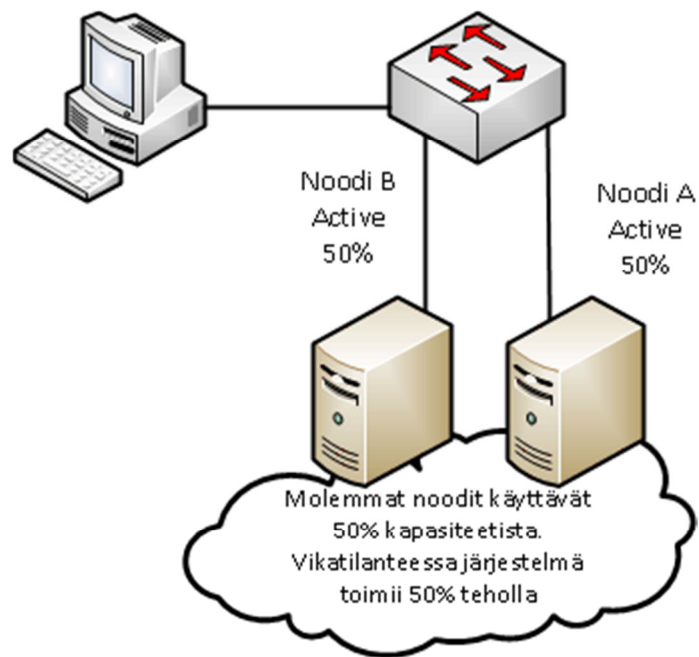
Termiä klusteri käytetään yleisesti kuvaamaan joukkoa hajautettuja prosessointijärjestelmiä, mutta termiä tietotekniikassa käyttävät eivät ole päässeet yhteisymmärrykseen määritelmästä. Gregory Pfister täytti yli 500 sivua kirjassaan *In Search of Clusters* (1997) pohtiessaan tätä ongelmaa. Hän ehdottaa seuraavaa eleganttia määritelmää. Klusteri on rinnakkainen tai hajautettu järjestelmä, joka sisältää kokoelman yhteenliitettyjä kokonaisia tietokoneita ja, jota käytetään yhtenä yhdistettynä laskentaresurssina. (Kopper 2005, 1.)

Tässä yhteydessä klusterilla tarkoitetaan ryhmää palvelimia, jotka toimivat yhdessä vikasietoisuuden parantamiseksi. Klusteri näkyy käyttäjille yhtenä resurssina. Jos itsenäinen palvelin vikaantuu, sen resursseja ruvetaan ajamaan toisessa palvelimessa. Itsenäisiä palvelimia kutsutaan klusterinoodiksi. Noodien tulisi myös olla vikasietoisia. Tavasta, jossa klusteri rakennetaan vikasietoista noodeista, voidaan käyttää sanontaa ”fault tolerance in depth”. (Snedaker 2004, 190-191.)

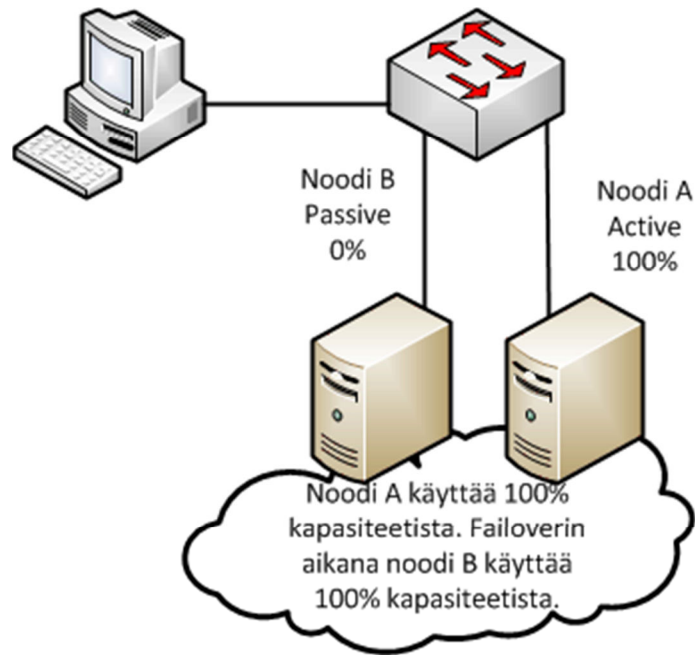
Windows Server 2003 -klusterit koostuvat useasta ohjelmasta, jotka kuuluvat kahteen kategoriaan. Ensimmäinen on klusterointiohjelmat (Clustering software), joka sisältää tarvittavat ohjelmat, jotta klusteri pystyisi toimimaan. Tärkeimmät näistä ohjelmista ovat resurssimonitori (Resource Monitor) ja klusteripalvelu (Cluster Service). Toinen kategoria on hallintaohjelmat (Administrative software), joilla hallitaan klusteria. (Ivens 2003, 812.) Kuvioissa 1, 2 ja 3 on esimerkkejä klustereiden toteutuksista.



Kuvio 1. Esimerkki klusteroidusta ratkaisusta (Shimonski 2003, 5).



Kuvio 2. Esimerkki Active/Active klusteroinnista (Shimonski 2003, 14).



Kuvio 3. Esimerkki Active/Passive klusteroinnista (Shimonski 2003, 13).

2.1 Klusteripalvelu

Klusteripalvelua ajetaan jokaisessa klusterinoodissa. Se kontrolloi kaikkea klusteritoimintaa. Tarkennettuna klusteripalvelu kontrolloi seuraavia toimintoja:

- kaikkien klusteriobjektien hallinnointi mukaan lukien noodit, jaetut tallennuslaitteet ja konfiguraatiot
- yhteyksien koordinointi muiden klusteripalveluiden kanssa klusterissa
- sovellus- ja noodivirheiden havaitseminen ja failover operaatioiden aloittaminen
- tapahtumien tiedotteiden hallinnointi. (Ivens 2003, 812.)

2.2 Resurssimonitori

Jokainen fyysinen tai looginen komponentti, joka on osa klusteria, on resurssi. Resurssimonitorit ovat ohjelmia, jotka toimivat ikään kuin yhteyshenkilönä klusteripalvelun ja resurssien välillä. Resurssimonitori mahdollistaa näiden resurssien ajon erillään toisistaan ja klusteripalvelusta. Tämä tukee klusterin luotettavuutta, saatavuutta ja skaalautuvuutta. Esimerkiksi resurssin

vikaantuminen ei vaikuta klusteripalvelun toimintaan. Resurssimonitori valvoo myös klusteripalvelua joten, jos se vikaantuu, resurssimonitori voi reagoida asettamalla kaikki noodin resurssit offline-tilaan. (Ivens 2003, 813.)

2.3 Loogiset komponentit

Virtual server

Virtual server koostuu palvelimen nimestä ja resurssiryhmistä, joita tarvitaan klusteroidun sovelluksen ajamiseen. Virtual server on clientin näkymä klusterista. Clientit eivät siis näe, mikä noodi ylläpitää virtual serveriä. (Ivens 2003, 814.)

Resurssit

Resurssi on klusterin pienin mahdollinen yksikkö. Resurssi voi olla fyysinen tai looginen, ja sitä voidaan hallinnoida yksittäin ja aktivoida tai deaktivoida itsenäisesti. Resurssin voi omistaa vain yksi noodi kerrallaan. Resurssin ominaisuudet määrittävät noodit, jotka saavat aktivoida resurssin (Ivens 2003, 815). Resursseja ovat esimerkiksi fyysiset levyasemat ja IP-osoite. (Snedaker 2004, 191.)

Ryhmät

Ryhmä on kokoelma resursseja. Jokainen resurssi voi kuulua vain yhteen ryhmään kerrallaan. Failoverin aikana klusteripalvelu siirtää ryhmän noodista toiseen varmistaen näin, että ryhmän kaikki resurssit ovat deaktivoituna ennen kuin ne aktivoidaan toisessa noodissa. Ryhmään kohdistunut operaatio vaikuttaa kaikkiin siihen kuuluviin resursseihin. (Ivens 2003, 815.)

2.4 Fyysiset komponentit

Noodit

Palvelin, joka kuuluu klusteriin, on noodi. Noodi on online-tilassa, kun palvelin on päällä, siinä ajetaan klusteripalvelua ja noodi kommunikoi muiden online-tilassa olevien noodien kanssa. (Ivens 2003, 815.) Noodeille voidaan

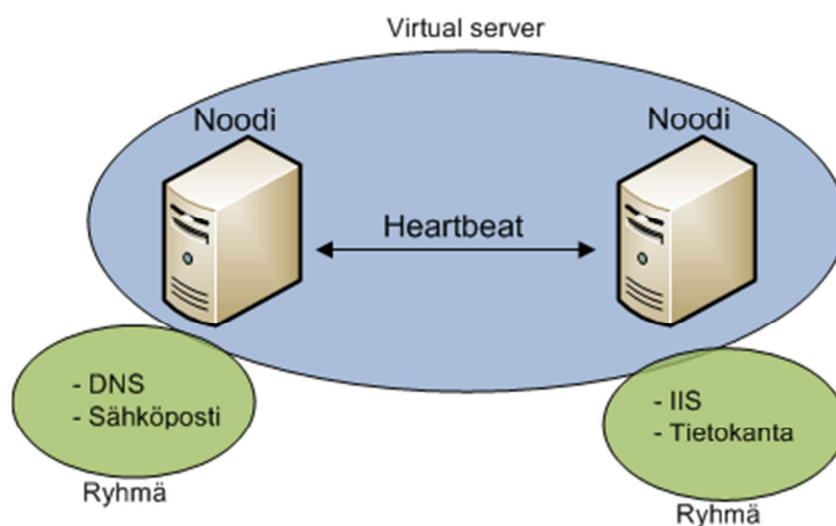
konfiguroida ensisijainen rooli, joka parantaa klusterin käyttöastetta, jos kaikki noodit ovat aktiivisia (Snedaker 2004, 191).

Klusterin levyasemat ja Quorum-resurssi

Klusterin levyasemat ovat loogisia levyasemia (joko yksittäisiä fyysisiä levyjä tai RAID-pakkoja), joihin kaikilla noodeilla on pääsy. Levyasemilla säilytetään sovellusten dataa. Quorum-resurssi on jaetulla väylällä oleva levyasema, joka ei kuulu mihinkään failover-ryhmään. Klusteripalvelu tallentaa hallintadatan ja palautuslokin quorum-resurssiin, ja välittää sen avulla klusterin hallintaa. (Ivens 2003, 815.)

Heartbeat

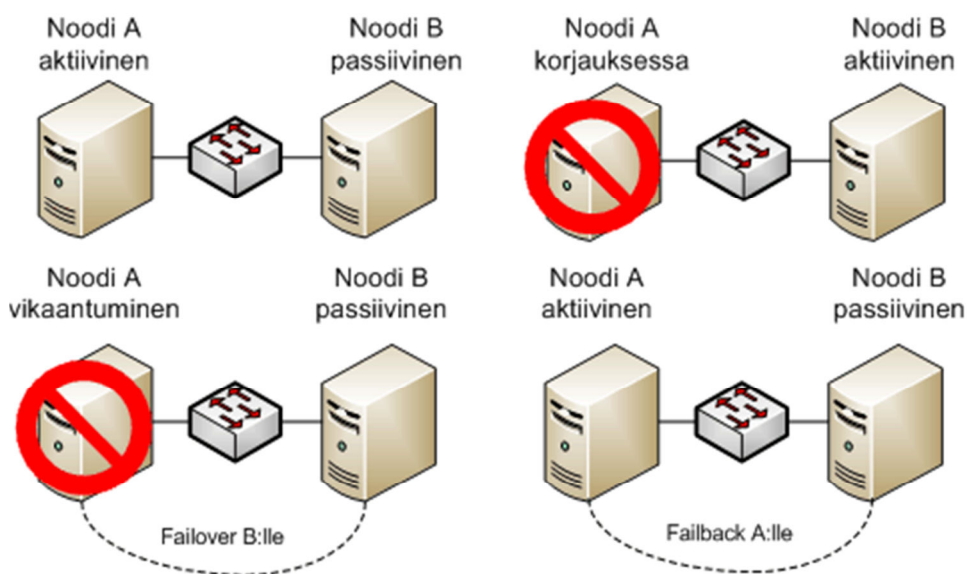
Heartbeatit ovat IP-paketteja, joita noodit lähettävät säännöllisesti. Kun noodi ei vastaanota heartbeat-pakettia, se olettaa, että noodi, joka ei lähettänyt pakettia on kaatunut. Tällaisessa tilanteessa aktiivinen noodi käynnistää failover-prosessin. (Ivens 2003, 815.) Kuviossa 4 on kuvattu klusterin eri komponentteja ja niiden vaikutusalueita.



Kuvio 4. Klusterin komponentteja.

2.5 Failover ja failback

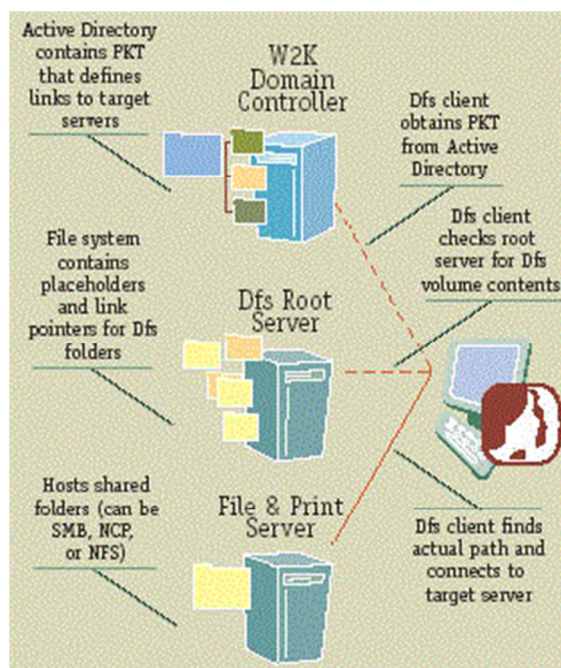
Failover on prosessi, joka käynnistyy, kun noodi vikaantuu. Vikaantunut noodi lopettaa resurssien suorittamisen ja niiden suorittaminen aloitetaan toisessa noodissa. Failoveria seuraa failback, kun vikaantunut noodi on korjattu. Näin noodiin palautetaan suoritettavaksi resurssit, joiden suorittaminen lopetettiin vikaantumisen takia. (Snedaker 2004, 192.) Kuviossa 5 on kuvattu failover- ja failback -prosessit.



Kuvio 5. Failover ja failback (Shimonski 2003, 11).

3 Distributed File System

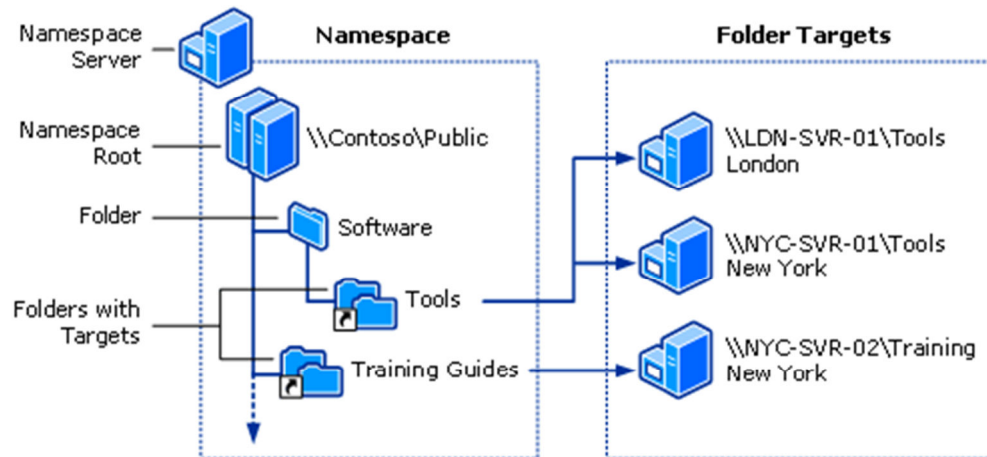
Distributed File System (DFS) on teknologia, minkä avulla voidaan yhdistää eri palvelimien jaetut hakemistot yhdeksi tai useammaksi ryhmäksi. Ryhmiä kutsutaan DFS-nimiavaruuksiksi. DFS-nimiavaruus on virtuaalinen näkymä jaetuista hakemistoista ja jossa käyttäjät navigoivat ilman tarvetta tietää tiedon varsinaista tallennuspaikkaa. (Microsoft 2003a.) Kuviossa 6 on kuvattu prosessia tiedon etsimisestä DFS:stä clientin näkökulmasta.



Kuvio 6. Prosessi tiedon etsimiseen DFS:stä clientin näkökulmasta (Boswell 2002).

3.1 Termit

DFS:ään liittyy joitain termejä, jotka tulisi ymmärtää. Nimiavaruuden alkupistettä kutsutaan DFS-juureksi (root). Juuri yhdistyy yhteen tai useampaan root targetiin, joista kukin vastaa jaettua kansiota erillisellä palvelimella. DFS-polku (path) on UNC (Universal Naming Convention) -polku, joka alkaa DFS-juuresta. DFS-polulla olevia osia kutsutaan DFS-linkeiksi, jotka yhdistävät yhden tai useamman link targetin. Link target voi olla mikä tahansa UNC-polku, esimerkiksi jaettu hakemisto tai toinen DFS-polku. (Microsoft 2003a.) Kuviossa 7 on esimerkki nimiavaruudesta.

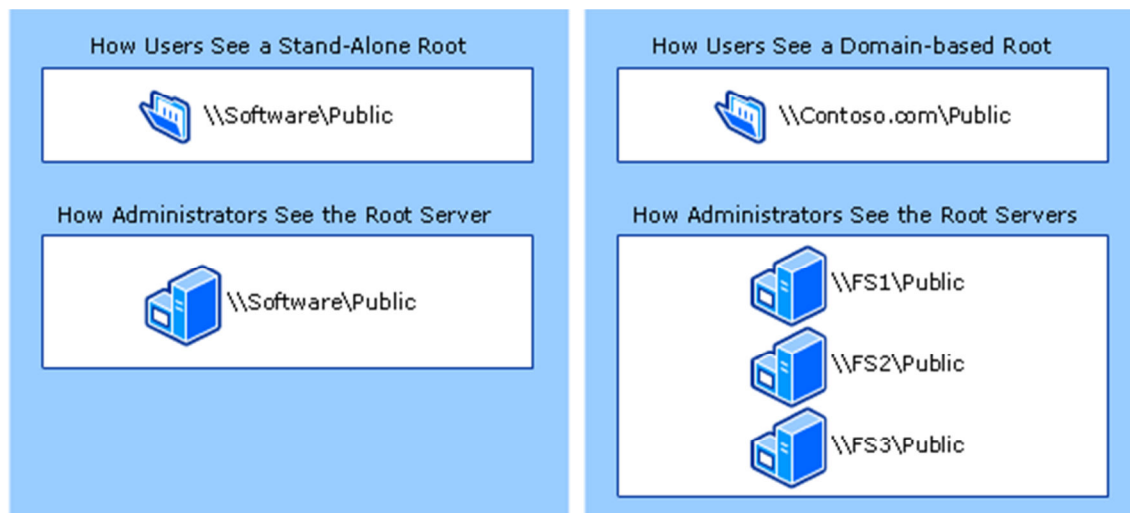


Kuvio 7. DFS nimiavaruus (Microsoft 2010).

3.2 Stand-alone ja domain-pohjainen DFS

DFS:n toteutukseen on kaksi ratkaisua, stand-alone ja domain-pohjainen. Stand-alone-toteutuksessa konfiguraatio varastoidaan paikallisesti palvelimen rekisterissä ja polku tai linkki alkaa isäntäpalvelimen nimellä. Stand-alone DFS-juurella on vain yksi root target. Stand-alone toteutukset eivät ole vikasietoisia. Jos root targetiin ei pääse, koko DFS-nimiavaruus on poissa käytöstä. Stand-alone DFS rooteista voi tehdä vikasietoisen klustereiden avulla. (Microsoft 2003a.)

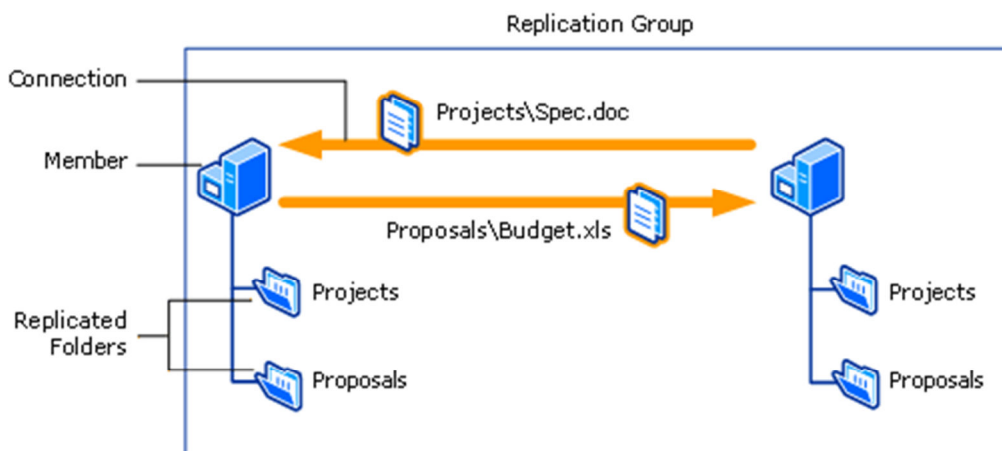
Domain-pohjaisessa DFS-toteutuksessa konfiguraatio on Active Directoryssa ja linkki tai polku alkaa domainnimellä. Domain DFS-juurella voi olla useita root targetteja, minkä vuoksi toteutus tarjoaa mahdollisuuden vikasietoisuuteen ja load balancingiin. (Microsoft 2003a.) Kuviossa 8 on esimerkki stand-alone ja domain pohjaisesta DFS-toteutuksesta.



Kuvio 8. Stand-alone ja domain-pohjainen DFS (Microsoft 2003a).

3.3 DFS-replikointi

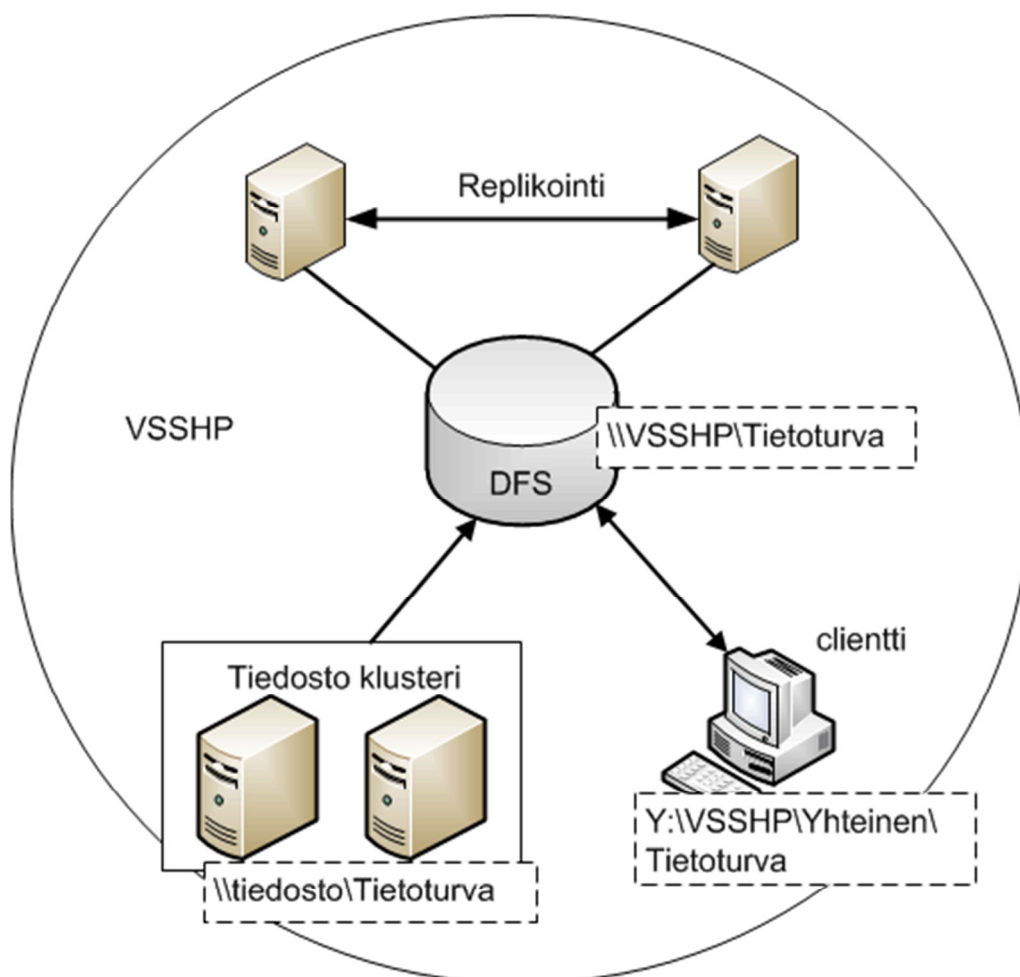
DFS-replikointia käytetään pitämään kansiot synkronoituina useammalla palvelimella. Datan replikointi usealle palvelimelle lisää saatavuutta ja antaa käyttäjille eri toimipisteissä nopean ja luotettavan pääsyn tiedostoihin. DFS-replikointiprotokolla käyttää remote differential compression (RDC) algoritmia, joka mahdollistaa tehokkaan tiedostojen päivityksen verkossa, jossa on rajallinen määrä kaistaa. RDC tunnistaa lisäykset, poistot ja muutokset tiedostoissa mahdollistaen vain muutosten replikoinnin, kun tiedostoja päivitetään. (Microsoft 2005.) Kuviossa 9 on esimerkki DFS-replikoinnista.



Kuvio 9. DFS-replikointi (Microsoft 2010).

3.4 DFS:n rajoitukset

On tärkeää olla tietoinen DFS-objektin koosta. Yksi merkintä metadatatassa vie noin 400 tavua DFS-linkkiä kohden, jolloin se ei sisällä kommentteja, joilla kuvataan DFS-nimiavaruutta. Keskikokoinen DFS-nimiavaruus, joka sisältää 100 linkkiä, vaatii noin 40 kilotavua (kB) tilaa. Mikä tahansa muutos nimiavaruudessa johtaa DFS-objektin replikointiin kaikkiin domain controllereihin. On suositeltavaa, että DFS-objekti domain pohjaiselle nimiavaruudelle olisi alle 5 megatavua (MB), jottei nimiavaruusmuutoksista johtuva liikenne kuormittaisi verkkoa liikaa. 5 MB:n nimiavaruus vastaa noin 5000 linkkiä. (Microsoft 2003a.) Kuviossa 10 on kuvattu VSSHP:n DFS:n rakenne.



Kuvio 10. VSSHP:n DFS:n rakenne.

4 Pääsynvalvontalistat

Pääsynvalvontalistoilla (ACL, Access control list) rajoitetaan käyttäjien pääsyä verkon eri osiin, palveluihin tai tämän tutkimuksen tapauksessa hakemistojakoihin. Riippuen siitä, mitä halutaan rajoittaa, ACL:t voivat sijaita infrastruktuurin eri laitteissa kuten reitittimissä ja palomureissa. Hakemistojakoihin pääsyä rajoitetaan yleensä levyrajapinnassa antamalla tiettyjä oikeuksia käyttäjille tai ryhmille.

Kaikilla objekteilla on myös omistaja, joka hallinnoi oikeuksia ja jakaa niitä käyttäjille. Windows Serverissä objektin luoja on oletuksena sen omistaja. Jos luoja kuuluu Administrators-ryhmään, ryhmä asetetaan objektin omistajaksi. Omistaja voi aina muuttaa oikeuksia, vaikka siltä olisi kielletty pääsy objektiin. (Microsoft 2003b.)

5 Skriptit

Skriptejä voidaan käyttää eri tavoin esimerkiksi konfigurointiin tai tiedon keruuseen. Hakemistojakojen kartoituksessa käytin skriptejä hakemistopuun ja oikeuksien kartoitukseen. Yksinkertaisella batch-tiedostolla sain hakemiston sisällön dir-komennolla ja pääsynvalvontalistan (ACL) Microsoftin XCACLS skriptillä, saatavilla osoitteessa <http://support.microsoft.com/kb/825751>. Skriptissä ei ollut muuttujia vaan jokaiselle kohteelle oli oma rivinsä. Skriptin luominen Excelissä merkkijonoja yhdistämällä oli silti parempi ratkaisu kuin komentojen ajaminen erikseen. Komennoista otettiin tekstiulostulo, mikä tuotiin Excel-asiakirjaan.

Kuvasta 1 näkee periaatteen, jota käytettiin skriptien luomisessa. Työkalu ei ole kovin monipuolinen ja vaatii käsisäättöjä, mutta ajoi tässä tapauksessa asiansa. Kuvassa on luotu ACL:t keräävä osa skriptistä.

F1		fx =-\$B\$9&A1&\$C\$9&B12&\$E\$9													
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Euron käyttöönotto					cscript xcaccls.vbs %y:\										!Euron käyttöönotto!"" > "%y_kansiot\scriptoutput\Euron kayttoonotto.txt"
Hallintotoimisto					cscript xcaccls.vbs %y:\										!Hallintotoimisto!"" > "%y_kansiot\scriptoutput\Hallintotoimisto.txt"
Kehittämistöimisto					cscript xcaccls.vbs %y:\										!Kehittämistöimisto!"" > "%y_kansiot\scriptoutput\Kehittämistöimisto.txt"
Mittarointi					cscript xcaccls.vbs %y:\										!Mittarointi!"" > "%y_kansiot\scriptoutput\Mittarointi.txt"
Ovk0402					cscript xcaccls.vbs %y:\										!Ovk0402!"" > "%y_kansiot\scriptoutput\Ovk0402.txt"
ovk042002					cscript xcaccls.vbs %y:\										!ovk042002!"" > "%y_kansiot\scriptoutput\ovk042002.txt"
Ovk0802					cscript xcaccls.vbs %y:\										!Ovk0802!"" > "%y_kansiot\scriptoutput\Ovk0802.txt"
ovk42002					cscript xcaccls.vbs %y:\										!ovk42002!"" > "%y_kansiot\scriptoutput\ovk42002.txt"
sql6backup	cscript xccl "" > "%y_kansiot\script.txt"				cscript xcaccls.vbs %y:\										!sql6backup!"" > "%y_kansiot\scriptoutput\sql6backup.txt"
Talsu02					cscript xcaccls.vbs %y:\										!Talsu02!"" > "%y_kansiot\scriptoutput\Talsu02.txt"
Talsu03					cscript xcaccls.vbs %y:\										!Talsu03!"" > "%y_kansiot\scriptoutput\Talsu03.txt"
Talsu3	Euron käyttöönotto				cscript xcaccls.vbs %y:\										!Talsu3!"" > "%y_kansiot\scriptoutput\Talsu3.txt"
Tarjous	Hallintotoimisto				cscript xcaccls.vbs %y:\										!Tarjous!"" > "%y_kansiot\scriptoutput\Tarjous.txt"
TP 2002	Kehittämistöimisto				cscript xcaccls.vbs %y:\										!TP 2002!"" > "%y_kansiot\scriptoutput\TP 2002.txt"
TP01	Mittarointi				cscript xcaccls.vbs %y:\										!TP01!"" > "%y_kansiot\scriptoutput\TP01.txt"
TP2001	Ovk0402				cscript xcaccls.vbs %y:\										!TP2001!"" > "%y_kansiot\scriptoutput\TP2001.txt"
tv	ovk042002				cscript xcaccls.vbs %y:\										!tv!"" > "%y_kansiot\scriptoutput\tv.txt"
Työsuojelu	Ovk0802				cscript xcaccls.vbs %y:\										!Työsuojelu!"" > "%y_kansiot\scriptoutput\Työsuojelu.txt"
Viestintä	ovk42002				cscript xcaccls.vbs %y:\										!Viestintäyksikkö!"" > "%y_kansiot\scriptoutput\Viestintäyksikkö.txt"

Kuva 1. Kuvakaappaus Excel-taulukosta, jossa luotiin skriptejä.

6 Tutkimuksen ongelmat

Tutkimuksen tekemisen suurimmat ongelmat olivat kohdehakemiston koko. Y-asemasta oli olemassa teknistä dokumentaatioita, mutta sen sisällöstä ei ollut olemassa aiempaa dokumentaatiota. Puuttuvat käytösäännöt ja valvonta olivat johtaneet Y-aseman muuttumiseen dokumenttien kaatopaikaksi. Quotan ja ylläpidon puuttuminen johtaa väistämättä asiattoman tai vanhan tiedon kerääntymiseen. Y-aseman tapauksessa suurempi ongelma on ollut vanhan tiedon kerääntyminen.

Ongelmia esiintyi myös dokumenttien luokittelussa. Vanhaa tietoa esiintyi lähes jokaisessa hakemistossa, mutta kaikki vanha tieto ei ollut välttämättä vanhentunutta. Luokittelussa etsittiin myös kansioita, jotka olivat tyhjiä. Tämä johti ongelmiin niiden kansioiden kohdalla, joiden tyhjiys saattoi johtua ajantasaisen tiedon puutteesta. Sekä vanhan tiedon että tyhjien kansioiden kohdalla mietittiin tarkasti, mihin luokkaan kansiot kuuluvat. Tiedon luokittelussa muodostui ongelmaksi myös sellainen sisältö, joka sopi useampaan luokkaan. Tutkimuksessa päätettiin luokitella tällainen sisältö sen pääasialliseen luokkaan. Näin ollen esimerkiksi vanhat sovellukset luokiteltiin sovelluksiksi ja vanhat muistiot ja pöytäkirjat kokousasiakirjoiksi eikä vanhaksi tiedoksi.

7 Hakemiston esitys

Hakemistoista kerätyn tiedon esitykseen valittiin Excel. Taulukossa olevaa tietoa on helppo käsitellä ja lukea. Tekstitiedostojen tuonti työkirjaan on helppoa ja tiedot tulevat automaattisesti taulukkomuotoon. Hakemiston sisältö kirjattiin omaan sarakkeeseen ja oikeudet esitettiin kommentteilla. Kommenttiruutuun liitettiin hakemiston pääsynvalvontalista (ACL). Koska usealla hakemistolla voi olla identtinen ACL, määritettiin oikeudet-sarakkeen otsikkosoluun kommentilla useimmiten käytössä ollut ACL. Jos hakemistolla oli tämä ACL, oikeudet-sarakkeeseen tuli merkintä default. Hakemistoilla saattoi olla myös eritysoikeuksia (special permissions), jotka näkyvät ACL:ssä kirjain- ja numeroyhdistelminä. Näiden yhdistelmien merkitys kirjattiin taulukkoon erikseen. Kuvissa 2 ja 3 on esimerkki taulukosta.

		hakemiston sisältö
<KANSIO>	.	
<KANSIO>	..	
<KANSIO>	Muut	vanhoja dokumentteja(2003
<KANSIO>	POSTOP	postoperatiivisia ohjeita (20
<KANSIO>	Postopvanha	vanhoja postop.ohjeita (200
<KANSIO>	Tilastot	tilastoja (2006 ja vanhempia)
<KANSIO>	Vuosikert	vanhoja vuosikertomuksia (1

Kuva 2. Esimerkki taulukosta, hakemiston sisältö.

oikeudet	Permissions:			
	Type	Username	Permissions	Inheritance
	Allowed	VSSHP\SgALEikkausFull	Full Control	This Folder, Subfolders and Files
	Allowed	VSSHP\Domain Users	Special (E6)	Files Only
default	Allowed	VSSHP\Domain Users	List Folder Contents	This Folder and Subfolders
default	Allowed	VSSHP\	Full Control	This Folder, Subfolders and Files
default	Allowed	VSSHP\	Full Control	This Folder, Subfolders and Files
default	No Auditing set			
default	Owner: BUILTIN\Järjestelmänvalvojat			

Kuva 3. Esimerkki taulukosta, hakemiston oikeudet.

8 Y-aseman vaikutus työntekijään

Nykyisellään hakemistojaon vaikutus työntekijään on huomattava. Asemaa käytetään muun muassa sanelujen väliaikaissäilytykseen, mikä on merkittävä osa sairaanhoitoa. Organisaatiosta riippuen asemaa voidaan käyttää myös tiedotukseen ja ohjeiden säilytykseen. Hakemistojaossa on myös paljon sovelluksia, Access-kantoja ja Excel-sovelluksia, joista hyvin harva, jos mikään, on koko käyttäjäkunnan käytössä. Vaikutus työntekijän elinkaareen vaihtelee siis organisaation mukaan, mikä vaikeuttaa hakemistojaon uudelleenorganisointia.

Y-aseamalla on kuitenkin hyvin laaja-alaisesti tietoa ja todennäköisesti jokainen työntekijä joutuu sitä käyttämään jossain vaiheessa. Joillain organisaatioilla voi olla myös muita levyjakoja, jotka sijaitsevat oikeasti Y-aseamalla. Myös esimerkiksi VSSHP:n intranetissä on dokumentteja, jotka sijaitsevat Y-aseamalla. Y-asema on siis tärkeä työntekijän kannalta riippumatta siitä, mihin organisaatioon hän kuuluu.

9 Tiedon luokittelu

Suurin osa Y-aseamalla sijaitsevasta datasta on dokumentteja. Dokumentit voivat olla esimerkiksi ohjeita, lomakkeita tai tiedotteita. Nämä kolme tyyppiä valittiin myös alaluokiksi dokumenttien luokittelussa niiden yleisyyden ja käytettävyyden vuoksi. Ne kaikki ovat olennainen osa hoitotyötä. Muita suuren esiintyvyyden omaavia tyyppejä ovat esimerkiksi kokousmuistiot ja -pöytäkirjat.

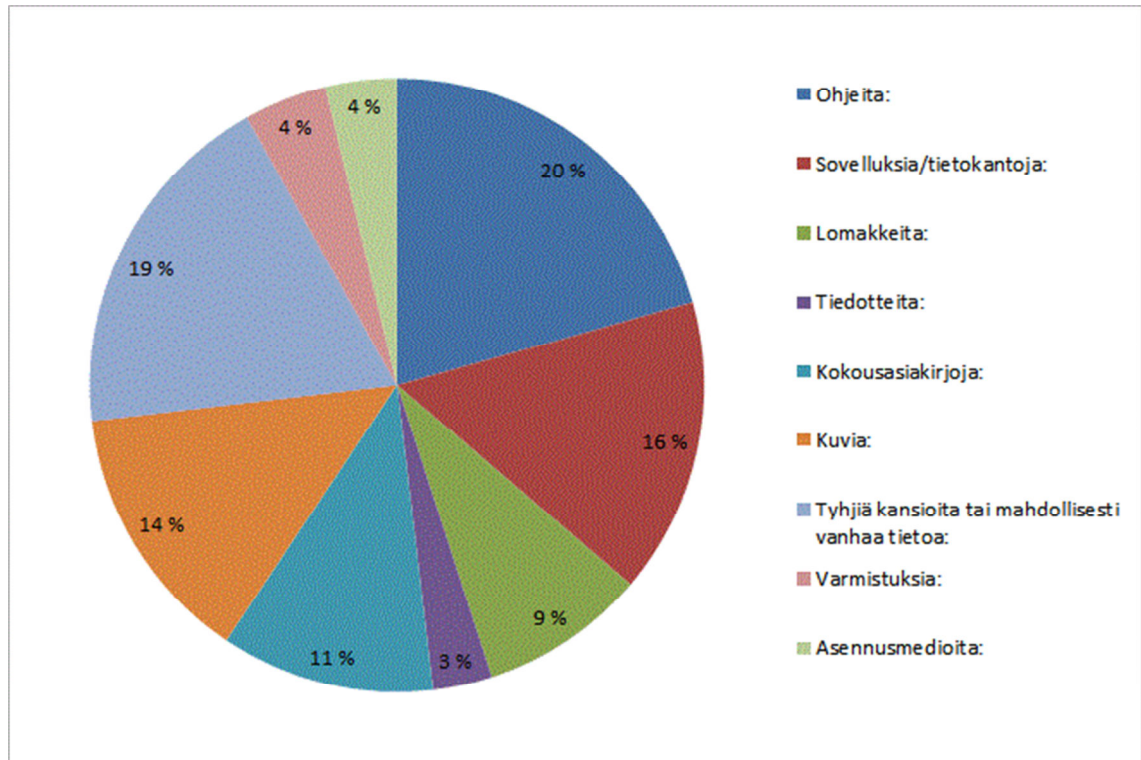
Yksi tutkimuksen pääkohdista oli kartoittaa sovellukset Y-aseamalla. Luonnollisesti siis niin kutsuttu aktiivinen sisältö eli sovellukset ja tietokannat oli keskeisessä osassa luokittelua. Koska kansioita on valtava määrä, ei jokaista ohjetta, muistiota tai lomaketta sisältävää kansiota kirjattu vaan luokittelu tehtiin pääosin siten, että kirjattiin, mitkä kansiot sisälsivät alikansioita, joissa oli olennaista tietoa.

Luokittelu koostui seuraavista osista

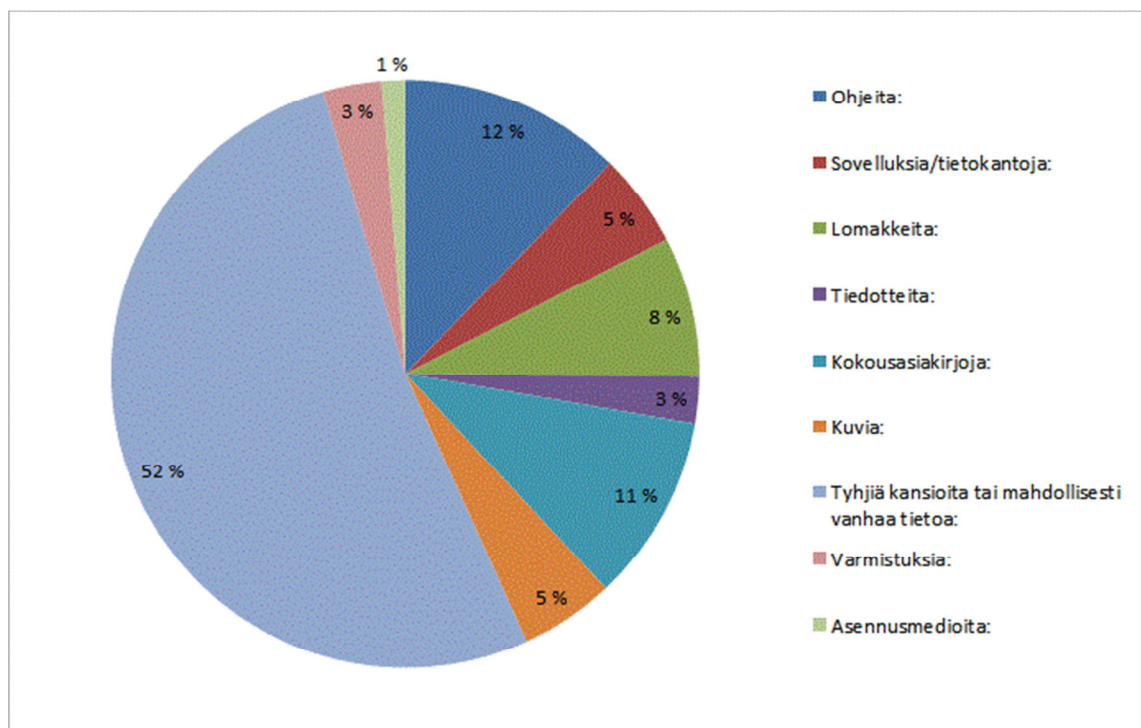
- ohjeet
- sovellukset/tietokannat
- lomakkeet
- tiedotteet
- kokousasiakirjat
- kuvat
- tyhjät kansiot tai mahdollisesti vanhaa tietoa
- varmistukset
- asennusmediat.

Kuvat-luokka ei sisällä potilaskuvia, logoja tai karttoja. Asennusmedioihin kuuluvat levykuvat ja sovellusten asennustiedostot. Näiden hallinnointi tulee tulevaisuudessa kuulumaan yksinomaan Medbitille ja asennukset on ulkoistettu Fujitsulle, joten muilla organisaatiolla ei ole tarvetta tähän luokkaan kuuluville tiedostoille (Heinonen 5.10.2010, sähköpostiviesti). Joidenkin organisaatioiden jäsenillä on ollut tavallista enemmän oikeuksia, joten asennusmedioita on kertynyt joihinkin hakemistoihin. Tästä syystä niiden kartoittaminen on tärkeää.

Seuraavista kuvioista 11 ja 12 selviävät eri luokkien osuudet luokitellusta sisällöstä TYKS-kansiossa ja muiden organisaatioiden yhteenlasketuista osuuksista. Kaaviot ovat vain suuntaa-antavia ja ne on laadittu tiedoista, mitkä kansiot sisältävät kyseisiin luokkiin kuuluvaa dataa, ei yksittäisistä tiedostoista.



Kuvio 11. Luokiteltu sisältö TYKS-kansiossa.



Kuvio 12. Luokiteltu sisältö muissa kansioissa.

10 Hakemiston uudelleenrakennus

Y-aseman uudelleenrakennuksessa tulisi kiinnittää erityistä huomioita siihen, että uudesta jaosta ei tule samanlaista kaatopaikkaa kuin vanhasta. Ennaltaehkäisyn työkaluja ovat quotat, käytösääntöjen luominen ja ylläpidon tehostaminen. Suurin ongelma käyttäjille tulee olemaan uuden työtavan omaksuminen. Aikaisemmin Y-asemalle on saanut tallentaa mitä tahansa eikä ole tarvinnut huolehtia vanhan tiedon poistamisesta, joten muutosvastarinta on väistämätöntä.

Quotat tulevat myös aiheuttamaan erimielisyyksiä johtuen organisaatioiden erilaisista tarpeista. Ennen niiden käyttöönottoa tulisi mahdollisia rajoituksia tarkastella käyttäjien näkökulmasta. Rajoitukset tuottaisivat lisää työtä myös ylläpidolle, koska joillekin organisaatiolle tulisi tarvetta lisätilalle. Muutosvastarinta aiheuttaisi todennäköisesti myös aiheettomia pyyntöjä lisätilasta. Pyyntöjen taustalla voi olla piittaamattomuus säännöistä, jolloin hakemistoon on kasautunut tarpeetonta tietoa tai jotain sinne kuulumatonta.

Quotaa tärkeämmässä asemassa ovat yhteiset käytösäännöt asemalle. Käyttäjille pitäisi tehdä selväksi, mihin asema on tarkoitettu. Sääntöjen noudattamisen valvonta on kuitenkin työlästä. Lisäksi eri organisaatiolla saattaa olla eri näkemyksiä yhteisen hakemiston käytöstä. Myös organisaatioiden todelliset tarpeet ovat erilaisia.

Osa vanhan Y-aseman kansiorakenteesta tulee todennäköisesti säilymään muutoksen jälkeen, ainakin ylemmillä tasoilla. Tämä ei ole ongelma vaan suorastaan suotavaa, jotta rakenne säilyisi selkeänä ja ehjänä. Organisaatiolla on myös samanlaisia tarpeita esimerkiksi ohjeiden muodossa. VSSHP:n aluesairaaloissa on samaa toimintaa kuin TYKSissä ja näissä organisaatioissa tehdään asiat ainakin hoitotyön osalta pääosin samalla tavalla. Ohjeita ei siis tarvitse olla kuin yhdessä paikassa, mihin jokaisella organisaation jäsenellä on pääsy. Vaikka dokumentit ovat suhteellisen pieniä kooltaan, päällekkäisyyksien poistaminen vähentää ylläpidon tarvetta.

Hakemistojaon uudelleenrakennuksessa tulisi ensiksi poistaa tarpeeton data. Koska tietohallintopalvelut eivät voi ottaa kantaa potilashoidollisiin asioihin, täytyisi hoitohenkilökunnan suorittaa käytännössä aseman siivous. Jotta siivous olisi mahdollisimman sujuvaa, voisi sen suorittaa hoitotyön ohessa siirtämällä data uuteen paikkaan silloin, kun sitä työssä tarvitsee. Tällainen prosessi vie tietysti huomattavasti aikaa ja muita resursseja, joten jossain vaiheessa täytyisi vanhan aseman käytöstä luopua ja siirtää sinne jäänyt data tarpeen vaatiessa. Koska kyseessä on näinkin suuri hakemistojako, jonka dokumentaatiota on lähes, ellei täysin mahdotonta ylläpitää, tarpeettoman datan siirtämistä tulisi välttää.

Hakemistojaon käyttö on hyvin moninaista ja uudelleenorganisoinnin aikana olisi hyvä mahdollisuus hajauttaa tieto usealle alustalle. Tämä selkeyttäisi myös hakemiston rakennetta. Tiedotuksen siirtyminen jollekin VSSHP:n yhteiselle alustalle olisi merkittävä uudistus myös siksi, että se tarjoaisi mahdollisuuden tiedotteiden parempaan hallintaan ja tiedon etsimiseen myös dokumenttien sisältä. Asemalla on myös huomattava määrä varsinaiseen hoitotyöhön liittymättömiä kuvia, joita varten voisi esimerkiksi perustaa VSSHP:n yhteisen kuvapankin. Potilaskuvien oikea paikka olisi luonnollisesti hoitokertomuksessa.

Hoitotyössä käytettävät sovellukset voisi siirtää omaan hakemistojakoon, jollainen on jo nyt käytössä. Kaikkien sovellusten säilytys samassa paikassa helpottaisi dokumentaation ja järjestelmien ylläpitoa.

Koska hoitotyössä ja sen hallinnoinnissa syntyy muutakin kuin hoitokertomuksia, täytyy tälle tiedolle olla myös jatkossa säilytyspaikka. Kaikelle tiedolle ei ole järkevää perustaa omia hakemistojakoja, joten Y-aseman kaltaiselle yleiselle hakemistojaolle on käyttöä ainakin lähitulevaisuudessa. Myös esimerkiksi SharePointissa olevan tiedon täytyy sijaita jossain.

Tulevaisuudessa väistämättä tulee eteen tilanne, jossa tallennuskapasiteetti loppuu. Vaikka fyysinen tallennustila ei tänä päivänä maksa juuri mitään, ei tallennuskapasiteetin lisäämisen pitäisi olla ensimmäinen ratkaisu tilan loppuessa. Väliaikaisen tiedon säilytysaika ja vanhentuneen tiedon poistaminen

ovat tärkeässä asemassa hakemistojaon ylläpidossa. Vastuulliseen liiketoimintaan kuuluu ympäristöystävällisyys ja levytilan lisäämisen vaikutuksia ympäristöön tulisi myös tarkastella.

11 Johtopäätökset

Tutkimuksen loppua kohden kävi selvemmäksi, että valitut työmenetelmät ja tapa, jolla tietoa esitettiin, eivät olleet parhaat mahdolliset. Pienemmässä mittakaavassa niillä olisi voitu saada järkeviä tuloksia, mutta ne eivät taipuneet näin suureen hakemistoon riittävän hyvin. Ilman olemassa olevaa dokumentaatiota tai mitään käsitystä siitä, mihin yhteistä hakemistojakoa käytetään, taso tasolta läpikäyminen on kuitenkin lähes ainoa vaihtoehto. Eri hakemistojen tasot eivät myöskään ole identtisiä määrässä ja käyttötarkoituksessa. Menetelmä ei siis ole ongelmaton.

Toisena vaihtoehtona on kyselyn teettäminen käyttäjillä, jonka tuloksena oltaisi tässä tutkimuksessa saatu tuhansia vastauksia, joiden läpikäynti olisi ollut vähintään yhtä työlästä. Ei voida olla myöskään täysin varmoja, että kyselyn tulokset olisivat kattaneet koko hakemiston. Realistisen kuvan saaminen edellyttäisi siis useampien menetelmien käyttöä ja näinkin isossa projektissa tekijöitä pitäisi olla useampia, jotta tuloksia syntyisi kohtuullisessa ajassa.

Dokumentaation pitäisi olla myös helposti luettavassa muodossa eikä useampaan taulukkoon hajautettu esitystapa palvele käyttäjää parhaalla mahdollisella tavalla. Taulukkoon mahtuu kuitenkin käytännössä suhteellisen vähän informaatioita ilman, että luettavuus kärsii.

LÄHTEET

Boswell, B. 2002. Simplify Resource Navigation With Dfs. Viitattu 17.9.2010 <http://redmondmag.com/articles/2002/08/01/simplify-resource-navigation-with-dfs.aspx>.

Ivens, K. 2003. Windows Server 2003: The Complete Reference. Emeryville, CA, USA: McGraw-Hill Osborne.

Kopper, K. 2005. The Linux Enterprise Cluster : build a highly available cluster with commodity hardware and free software. San Francisco, CA, USA: No Starch Press, Inc.

Microsoft 2003a. How DFS Works: Remote File Systems. Viitattu 19.7.2010 <http://technet.microsoft.com/en-us/library/cc782417%28WS.10%29.aspx>.

Microsoft 2003b. How Permissions Work. Viitattu 17.9.2010 <http://technet.microsoft.com/en-us/library/cc783530%28WS.10%29.aspx>.

Microsoft 2005. File Server Management Overview. Viitattu 1.11.2010 <http://technet.microsoft.com/en-us/library/cc758451%28WS.10%29.aspx>.

Microsoft 2010. DFS Step-by-Step Guide for Windows Server 2008. Viitattu 20.10.2010 <http://technet.microsoft.com/en-us/library/cc732863%28WS.10%29.aspx>.

Shimonski, R. 2003. Windows Server 2003 Clustering & Load Balancing. Emeryville, CA, USA: McGraw-Hill Osborne.

Snedaker, S. 2004. Best Damn Windows Server 2003 Book Period. Rockland, MA, USA: Syngress Publishing.